TMA 4115 Matematikk 3

Lecture 26 for MTFYMA

Alexander Schmeding

NTNU

18. April 2016

In today's lecture we will

- review the Gram-Schmidt algorithm
- apply it to study least-square problems
- solve some least square problems associated to linear models

Inner product, length and orthogonality

Let $\overrightarrow{x}, \overrightarrow{y}$ be vectors in $\mathbb{R}^n$ with components x_i and y_i, respectively.

Dot product/ Inner product $\overrightarrow{x} \cdot \overrightarrow{y} = \overrightarrow{x}^T \overrightarrow{y} = \sum_{i=1}^n x_i y_i$.

$\overrightarrow{x}$ and $\overrightarrow{y}$ are orthogonal to each other if and only if $\overrightarrow{x} \cdot \overrightarrow{y} = 0$.

For a subspace $W \subseteq \mathbb{R}^n$ we can split $\overrightarrow{y} \in \mathbb{R}^n$ uniquely into

$$\overrightarrow{y} = \overrightarrow{y}_W + \overrightarrow{z}, \quad \overrightarrow{y}_W \in W, \ \overrightarrow{z} \in W^\perp$$

and obtain an orthogonal projection $\text{proj}_W \colon \mathbb{R}^n \to W, \overrightarrow{y} \mapsto \overrightarrow{y}_W$.
If $\{ \overrightarrow{v}_1, \ldots, \overrightarrow{v}_p \}$ is an orthogonal basis for W, its easy to compute

$$\overrightarrow{y}_W = \sum_{i=1}^p \frac{\overrightarrow{y} \cdot \overrightarrow{v}_i}{\overrightarrow{v}_i \cdot \overrightarrow{v}_i} \overrightarrow{v}_i$$

21.7 The Gram-Schmidt Process

Let $W = \text{span}\{\overrightarrow{x}_1, \ldots, \overrightarrow{x}_p\} \subseteq \mathbb{R}^n$ and $\overrightarrow{x}_i \neq 0$ for $1 \leq i \leq p$. Define

$$\begin{aligned}
\overrightarrow{u}_1 &= \overrightarrow{x}_1 \\
\overrightarrow{u}_2 &= \overrightarrow{x}_2 - \frac{\overrightarrow{x}_2 \cdot \overrightarrow{u}_1}{\overrightarrow{u}_1 \cdot \overrightarrow{u}_1}\overrightarrow{u}_1 \quad \text{if } \overrightarrow{u}_2 = \overrightarrow{0} \text{ discard and construct} \\
& \qquad \qquad \text{with same formula, now for } \overrightarrow{x}_3 \\
\vdots &= \vdots \qquad \vdots \qquad \qquad \vdots \\
\overrightarrow{u}_p &= \overrightarrow{x}_p - \sum_{i=1}^{p-1} \frac{\overrightarrow{x}_p \cdot \overrightarrow{v}_i}{\overrightarrow{v}_i \cdot \overrightarrow{v}_i}\overrightarrow{v}_i \qquad \text{if } \overrightarrow{u}_p = \overrightarrow{0} \text{ discard}
\end{aligned}$$

Then $\{\overrightarrow{u}_1, \ldots, \overrightarrow{u}_d\}$ is an orthogonal basis for W (for some $1 \leq d \leq p$) and

$$\text{span}\,\{\overrightarrow{u}_1, \ldots, \overrightarrow{u}_k\} = \text{span}\,\{\overrightarrow{x}_1, \ldots, \overrightarrow{x}_k\} \quad \text{for } 1 \leq k \leq d$$

Application for Gram-Schmidt

(Spring 2011 Problem 5)

(a) Find an orthogonal basis of Col(A) with

$$A = \begin{bmatrix} 1 & 3 & 0 & 1 \\ 2 & 1 & 5 & -3 \\ -1 & -1 & -2 & 1 \end{bmatrix}$$

(b) Find for $\overrightarrow{y} = \begin{bmatrix} 1 \\ 7 \\ 3 \end{bmatrix}$ the nearest point in Col(A)

Let $\overrightarrow{x}_i$ be the i-th column of A. Recall Col(A) = span$\{\overrightarrow{x}_1, \overrightarrow{x}_2, \overrightarrow{x}_3, \overrightarrow{x}_4\}$.

Apply Gram-Schmidt to the generating system $\{\overrightarrow{x}_1, \overrightarrow{x}_2, \overrightarrow{x}_3, \overrightarrow{x}_4\}$ of Col(A).

Gram Schmidt on a generating system

Note that all $\vec{x}_i$ are non-zero. Thus set $\vec{u}_1 = \vec{x}_1 = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}$

Then $\vec{u}_2 = \vec{x}_2 - \frac{\vec{x}_2 \cdot \vec{u}_1}{\vec{u}_1 \cdot \vec{u}_1} \vec{u}_1 = \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix}$,

$\vec{u}_3 = \vec{x}_3 - \frac{\vec{x}_3 \cdot \vec{u}_1}{\vec{u}_1 \cdot \vec{u}_1} \vec{u}_1 - \frac{\vec{x}_3 \cdot \vec{u}_2}{\vec{u}_2 \cdot \vec{u}_2} \vec{u}_2 = \vec{0}$ we get $\vec{0}$!
$\rightsquigarrow$ discard $\vec{u}_3$ and continue with the next vector $\vec{x}_4$.

$\vec{u}_3 = \vec{x}_4 - \frac{\vec{x}_4 \cdot \vec{u}_1}{\vec{u}_1 \cdot \vec{u}_1} \vec{u}_1 - \frac{\vec{x}_4 \cdot \vec{u}_2}{\vec{u}_2 \cdot \vec{u}_2} \vec{u}_2 = \vec{0}$
again $\vec{0}$, we discard!Since it was the last, an orthogonal basis of Col(A) is $\left\{ \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}, \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix} \right\}$

Application of Gram-Schmidt

To answer part (b), we recall the

21.6 Best approximation theorem

$W \subseteq \mathbb{R}^n$ subspace, $\overrightarrow{y} \in \mathbb{R}^n$. Then $\overrightarrow{y}_W = \text{proj}_W(\overrightarrow{y})$ is closest in W to $\overrightarrow{y}$:

$$\|\overrightarrow{y} - \overrightarrow{y}_W\| < \|\overrightarrow{y} - \overrightarrow{w}\|, \quad \overrightarrow{w} \in W \setminus \{\overrightarrow{y}_W\}$$

We need to compute $\overrightarrow{y}_{\text{Col}(A)}$ as the nearest point to $\overrightarrow{y}$ in $\text{Col}(A)$.

$$\overrightarrow{y}_{\text{Col}(A)} = \frac{\overrightarrow{y} \cdot \overrightarrow{v}_1}{\overrightarrow{v}_1 \cdot \overrightarrow{v}_1} \overrightarrow{v}_1 + \frac{\overrightarrow{y} \cdot \overrightarrow{v}_2}{\overrightarrow{v}_2 \cdot \overrightarrow{v}_2} \overrightarrow{v}_2 = \begin{bmatrix} 0 \\ 5 \\ -2 \end{bmatrix}$$

Weight of baggage

What is the weight x of our suitcase?

If it is too heavy we have to pay extra on the airplane and we don't want that.
$\to$ try to get an accurate weight.

On a scale	Second scale	Third scale
		
weight $x = x_1$kg	weight $x = x_2$kg	weight $x = x_3$kg

What is the weight of our suitcase?

Weight of baggage II

From our scales we have the information:

$$\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \leftrightarrow \begin{cases} x & = x_1 \\ x & = x_2 \\ x & = x_3 \end{cases}$$

Unless $x_1 = x_2 = x_3$ (unlikely), the linear system is **inconsistent**, i.e. has no exact solution.

Problem: No solution!

Can we find an "approximate" solution? This means a number y such that the errors $|x_i - y|$ are as small as possible.

Weight of baggage III

Idea: If $A\vec{x} = \vec{b}$ is inconsistent, lets try to find $\vec{x}$ such that $A\vec{x}$ is as near as possible to $\vec{b}$.

Note $A\vec{x} \in \mathrm{Col}(A)$, hence if it is minimal $A\vec{x} = \mathrm{proj}_{\mathrm{Col}(A)}(\vec{b})$!

In our example $A = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ and $\vec{b} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$, we set $\vec{u}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$,
then $\mathrm{Col}(A) = \mathrm{span}\{\vec{u}_1\}$ and

$$\mathrm{proj}_{\mathrm{Col}(A)}(\vec{b}) = \frac{\vec{b} \cdot \vec{u}_1}{\vec{u}_1 \cdot \vec{u}_1}\vec{u_1} = \begin{bmatrix} \frac{x_1+x_2+x_3}{3} \\ \frac{x_1+x_2+x_3}{3} \\ \frac{x_1+x_2+x_3}{3} \end{bmatrix}$$

Thus $Ax = \mathrm{proj}_{\mathrm{Col}(A)}(\vec{b})$ shows: $x = \frac{x_1+x_2+x_3}{3}$.